

CLAIMS

What is claimed is:

We claim:

1. A method, comprising:

generating a list of one or more related documents ranked based upon relevance to a first representation of content associated with a first field of a reference extensible markup language document, the first representation including of a set of terms and one or more weighted values associated with each term in the set of terms; and

generating a link to each of the one or more related documents.

2. The method of claim 1, wherein the first field in the reference extensible markup language document is specified at the time a query is generated.

3. The method of claim 1, wherein the one or more related documents comprise a first related document having a second field, at the time the query is generated, a user specifies to search content associated with the second field.

4. The method of claim 1, wherein the reference extensible markup language document is selected from a group of documents in a database.

5. The method of claim 1, further comprising:

submitting the reference extensible markup language document to an engine for analysis.

6. The method of claim 1, wherein the link is a hypertext link.
7. The method of claim 3, wherein the second field of the related document contains semantically similar content to the content associated with the first field of the reference extensible markup language document.
8. The method of claim 1, further comprising:
executing a query on the reference extensible markup language document to generate the list and the link without a user having to request the query.
9. The method of claim 1, wherein the list further includes references to relevant fields within each related document.
10. The method of claim 1, wherein the set of terms includes singular terms and higher order terms.
11. The method of claim 1, wherein the set of terms includes singular terms and noun phrases.

12. The method of claim 1, wherein the set of terms includes higher order terms and proper names.

13. An apparatus, comprising:

a memory to store a first representation of content associated with one or more specified fields of a reference extensible markup language document, the first representation including a set of terms and one or more weighted values for each term in the set of terms; and

an engine having a reference extensible markup language document input and a specified field input, the engine to generate a list of one or more related documents and a link to each of the one or more related documents, the one or more related documents ranked based upon relevance to the first representation of content associated with the one or more specified fields of the reference extensible markup language document.

14. The apparatus of claim 13, further comprising:

a database of documents.

15. The apparatus of claim 13, further comprising:

a database containing a plurality of representations, each representation being associated with content in a particular field in an extensible markup language document.

16. The apparatus of claim 13, wherein the engine adjusts the one or more weighted values for each particular term in the set of terms by a comparison to a historical weighted value associated with each particular term in the set of terms.

17. The apparatus of claim 13, further comprising:

a converter to convert a non-extensible markup language document into an extensible markup language format.

18. The apparatus of claim 17, wherein the non-extensible markup language document is content associated with an e-mail, content associated with a web page, or content associated with a software application.

19. The apparatus of claim 13, wherein the engine has a module to compare the first representation to a plurality of representations in a database in order to identify documents that are most similar to the first representation.

20. The apparatus of claim 13, wherein the engine executes a query on the reference extensible markup language document to generate the list and the link without a user having to request the query.

21. A method, comprising:

receiving a reference extensible markup language document as a first input to an engine;

specifying a first field in the reference extensible markup language document as a second input to an engine;

generating a list of related documents ranked based upon their semantic similarity to content in the first field in the reference extensible markup language document; and generating a link to each related document in the list.

22. The method of claim 21, wherein the reference extensible markup language document has a first extensible markup language schema, and a first related extensible markup language document has a second extensible markup language schema.

23. The method of claim 21, further comprising:

identifying a first representation of content associated with the reference extensible markup language document, the first representation including a first set of terms and one or more weighted values associated with each term in the first set of terms; and

identifying a second representation of content associated with a second field in a first related extensible markup language document, the second representation including a second set of terms and a second weighted value associated with each term in the second set of terms.

24. An apparatus, comprising:

means for receiving a reference extensible markup language document as a first input to an engine;

means for generating a list of related documents ranked based upon their semantic similarity to content in the first field in the reference extensible markup language document; and

means for generating a link to each related document in the list.

25. The method of claim 24, wherein the reference extensible markup language document has a first extensible markup language schema, and a first related extensible markup language document has a second extensible markup language schema.

26. The method of claim 24, further comprising:

means for identifying a first representation of content associated with the reference extensible markup language document, the first representation including a first set of terms and one or more weighted values associated with each term in the first set of terms; and

means for identifying a second representation of content associated with a second field in a first related extensible markup language document, the second representation including a second set of terms and a second weighted value associated with each term in the second set of terms.

27. A machine-readable medium that provides instructions, which when executed by a machine, cause the machine to perform operations comprising:

generating a list of one or more related documents ranked based upon relevance to a first representation of content associated with a first field of a reference extensible

markup language document, the first representation including of a set of terms and one or more weighted values associated with each term in the set of terms; and generating a link to each of the one or more related documents.

28. The method of claim 27, further comprising:

executing a query on the reference extensible markup language document to generate the list and the link without a user having to request the query.
